

PCT09

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/936,823

DATE: 10/04/2001

TIME: 11:12:32

Input Set : A:\36838SEQ.txt

Output Set: N:\CRF3\10042001\I936823.raw

3 <110> APPLICANT: Valtion teknillinen tutkimuskeskus
 5 <120> TITLE OF INVENTION: Process for partitioning of molecules
 7 <130> FILE REFERENCE: 31805
 C--> 9 <140> CURRENT APPLICATION NUMBER: US/09/936,823
 C--> 10 <141> CURRENT FILING DATE: 2001-09-18
 12 <160> NUMBER OF SEQ ID NOS: 42
 14 <170> SOFTWARE: PatentIn Ver. 2.2
 16 <210> SEQ ID NO: 1
 17 <211> LENGTH: 428
 18 <212> TYPE: DNA
 19 <213> ORGANISM: Trichoderma reesei
 21 <220> FEATURE:
 22 <221> NAME/KEY: intron
 23 <222> LOCATION: (167)..(236)
 25 <220> FEATURE:
 26 <221> NAME/KEY: intron
 27 <222> LOCATION: (323)..(386)
 29 <220> FEATURE:
 30 <223> OTHER INFORMATION: Coding sequence of hfb1
 32 <400> SEQUENCE: 1
 33 atgaagttct tcgccatcgc cgctctcttt gccgccgctg ccgttgccca gcctctcgag 60
 34 gaccgcagca acggcaacgg caatgtttgc cctcccgcc tcttcagcaa cccccagtgc 120
 35 tgtgccacc aagtccttgg cctcatcggc cttgactgca aagtcgta gttgagccat 180
 36 aacataagaa tcctcttgac ggaaatatgc cttctcactc ctttaccct gaacagctc 240
 37 ccagaacgtt tacgacggca ccgacttccg caacgtctgc gccaaaaccg gcgccagcc 300
 38 tctctgctgc gtggcccccg ttgtaagttg atgccccagc tcaagctcca gtctttggca 360
 39 aaccattct gacaccaga ctgcaggccg gccaggctct tctgtgccag accgccgctc 420
 40 gtgcttga 428
 43 <210> SEQ ID NO: 2
 44 <211> LENGTH: 78
 45 <212> TYPE: DNA
 46 <213> ORGANISM: Artificial Sequence
 48 <220> FEATURE:
 49 <223> OTHER INFORMATION: Description of Artificial Sequence: PCR 5' primer
 51 <400> SEQUENCE: 2
 52 tcgggcacta cgtgccagta tagcaacgac tactactcgc aatgccttgt tccgcgtggc 60
 53 tctagttctg gaaccgca 78
 56 <210> SEQ ID NO: 3
 57 <211> LENGTH: 30
 58 <212> TYPE: DNA
 59 <213> ORGANISM: Artificial Sequence
 61 <220> FEATURE:
 62 <223> OTHER INFORMATION: Description of Artificial Sequence: PCR 3' primer
 64 <400> SEQUENCE: 3
 65 tcgtacggat cctcaagcac cgacggcggt 30
 68 <210> SEQ ID NO: 4
 69 <211> LENGTH: 63

ENTERED

P.5

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/936,823

DATE: 10/04/2001
 TIME: 11:12:32

Input Set : A:\36838SEQ.txt
 Output Set: N:\CRF3\10042001\I936823.raw

```

70 <212> TYPE: DNA
71 <213> ORGANISM: Artificial Sequence
73 <220> FEATURE:
74 <223> OTHER INFORMATION: Description of Artificial Sequence: PCR 5' primer
76 <400> SEQUENCE: 4
77   actacacgga ggagctcgac gacttcgagc agcccgagct gcacgcagag caacggcaac 60
78   ggc 63
81 <210> SEQ ID NO: 5
82 <211> LENGTH: 2211
83 <212> TYPE: DNA
84 <213> ORGANISM: Trichoderma reesei
86 <220> FEATURE:
87 <221> NAME/KEY: promoter
88 <222> LOCATION: (1)..(2211)
89 <223> OTHER INFORMATION: cbh1 promoter sequence
91 <400> SEQUENCE: 5
92   gaattctcac ggtgaatgta ggccttttgt agggtaggaa ttgtcactca agcacccecca 60
93   acctccatta cgctccccc atagagtcc caatcagtga gtcattggcac tgttctcaaa 120
94   tagattgggg agaagttgac ttccgcccag agctgaaggt cgcacaaccg catgatatag 180
95   ggtcggcaac ggcaaaaaag cacgtggctc accgaaaagc aagatgtttg cgatctaaca 240
96   tccaggaacc tggatacatc catcatcacg cagcaccact ttgatctgct ggtaaaactcg 300
97   tattcgccct aaaccgaagt gcgtggtaaa tctacacgtg ggcccctttc ggtatactgc 360
98   gtgtgtcttc tctaggtgca ttctttcctt cctctagtgt tgaattgttt gtgttgggag 420
99   tccgagctgt aactacctct gaatctctgg agaatggtgg actaacgact accgtgcacc 480
100  tgcattcatgt atataatagt gatcctgaga aggggggttt ggagcaatgt gggactttga 540
101  tggatcatcaa acaaagaacg aagacgcctc ttttgcaaag ttttgtttcg gctacgggtga 600
102  agaactggat acttgttgtg tcttctgtgt atttttgttg caacaagagg ccagagacaa 660
103  tctattcaaa caccaagctt gctcttttga gctacaagaa cctgtggggt atatatctag 720
104  agttgtgaag tcggtaatcc cgctgtatag taatacgagt cgcattctaa tactccgaag 780
105  ctgctgcgaa cccggagaat cgagatgtgc tggaaagctt ctacgcagcg gctaaattag 840
106  catgaaaggc tatgagaaat tctggagacg gcttgttgaa tcatggcgtt ccattcttcg 900
107  acaagcaaag cgttccgtcg cagtagcagg cactcattcc cgaaaaaact cggagattcc 960
108  taagtagcga tggaaaccga ataataaat aggcaataca ttgagttgcc tcgacggttg 1020
109  caatgcaggg gtactgagct tggacataac tgttccgtac cccacctctt ctcaaccttt 1080
110  ggcgtttccc tgattcagcg taccctgaca agtcgtaate actattaacc cagactgacc 1140
111  ggacgtgttt tgcccttcat ttggagaaat aatgtcattg cgatgtgtaa tttgcctgct 1200
112  tgaccgactg gggctgttcg aagcccgaat gtaggattgt tatccgaact ctgctcgtag 1260
113  aggcattgtg tgaatctgtg tcgggcagga cagcctcga aggttcacgg caagggaaac 1320
114  caccgatagc agtgtctagt agcaacctgt aaagccgcaa tgcagcatca ctggaaaata 1380
115  caaaccaatg gctaaaagta cataagttaa tgcctaaaga agtcatatac cagcggctaa 1440
116  taattgtaca atcaagtggc taaacgtacc gtaatttgcc aacgcgttgt ggggttgacg 1500
117  aagcaacggc aaagcccact tcccacgttt gtttcttcac tcagtccaat ctgagctggt 1560
118  gatcccccaa ttgggtcgct tgtttgttcc ggtgaagtga aagaagacag aggtagaagt 1620
119  gtctgactcg gagcgttttg catacaacca agggcagtga tggaaagacag tgaaatgttg 1680
120  acattcaagg agtatttagc cagggatgct tgagtgtatc gtgtaaggag gtttgtctgc 1740
121  cgatacgacg aatactgtat agtcacttct gatgaagtgg tccatattga aatgtaagtc 1800
122  ggcactgaac aggcaaaaaga ttgagttgaa actgcctaag atctcgggcc ctcgggcttc 1860
123  ggctttgggt gtacatgttt gtgctccggg caaatgcaaa gtgtggtagg atcgacacac 1920
124  tgctgccttt accaagcagc tgagggatat tgataggcaa atgttcaggg gccactgcat 1980

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/936,823

DATE: 10/04/2001

TIME: 11:12:32

Input Set : A:\36838SEQ.txt

Output Set: N:\CRF3\10042001\I936823.raw

```

125   ggtttcgaat agaaagagaa gcttagccaa gaacaatagc cgataaagat agcctcatta 2040
126   aacgaaatga gctagtaggc aaagtcagcg aatgtgtata tataaagggt cgagggtccgt 2100
127   gcctccctca tgctctcccc atctactcat caactcagat cctccaggag acttgtacac 2160
128   catcttttga ggcacagaaa cccaatagtc aaccgcggac tgcgcatcat g          2211
131 <210> SEQ ID NO: 6
132 <211> LENGTH: 1588
133 <212> TYPE: DNA
134 <213> ORGANISM: Trichoderma reesei
136 <220> FEATURE:
137 <223> OTHER INFORMATION: T. reesei eg11 cDNA
139 <400> SEQUENCE: 6
140   cccccctatc ttagtccttc ttgttgctcc aaaatggcgc cctcagttac actgccgttg 60
141   accacggcca tcttgcccat tgcccggctc gtcgccgccc agcaaccggg taccagcacc 120
142   cccgagggtc atcccaagtt gacaacctac aagtgtacaa agtcgggggg gtgcgtggcc 180
143   caggacacct cgggtggtcct tgactggaac taccgctgga tgcacgacgc aaactacaac 240
144   tcgtgcaccg tcaacggcgg cgtcaacacc acgctctgcc ctgacgaggc gacctgtggc 300
145   aagaactgct tcatcgaggg cgtcgactac gccgcctcgg gcgtcacgac ctcgggcagc 360
146   agcctcacca tgaaccagta catgcccagc agctctggcg gctacagcag cgtctctcct 420
147   cggctgtatc tcttggaact tgacggtgag tacgtgatgc tgaagctcaa cggccaggag 480
148   ctgagcttcg acgtcgacct ctctgctctg ccgtgtggag agaacggctc gctctacctg 540
149   tctcagatgg acgagaacgg gggcgccaac cagtataaca cggccgggtg caactacggg 600
150   agcggctact gcgatgctca gtgccccgtc cagacatgga ggaacggcac cctcaacact 660
151   agccaccagg gcttctgctg caacgagatg gatatactgg agggcaactc gagggcgaat 720
152   gccttgaccc ctcaactctt cacggccacg gcctgcgact ctgccgggtg cggcttcaac 780
153   ccctatggca gcggtacaa aagctactac ggccccggag ataccgttga cacctccaag 840
154   accttcacca tcatcaccca gttaaacacg gacaacggct cgcctcggg caaccttgtg 900
155   agcatcaccg gcaagtacca gcaaaacggc gtcgacatcc ccagcgccca gcccggcggc 960
156   gacaccatct cgtcctgccc gtccgcctca gcctacggcg gcctcgccac catgggcaag 1020
157   gccctgagca gcggcatggt gctcgtgttc agcatttggg acgacaacag ccagtacatg 1080
158   aactggctcg acagcggcaa cgccggcccc tgcagcagca ccgagggcaa cccatccaac 1140
159   atcctggcca acaaccccaa cagcgcgtc gtcttctcca acatccgctg gggagacatt 1200
160   gggctacta cgaactcgac tgcgcccccg ccccgccctg cgtccagcac gacgtttctg 1260
161   actacacgga ggagctcgac gacttcgagc agcccagact gcacgcagac tcaactgggg 1320
162   cagtgcgggt gcattgggta cagcgggtgc aagacgtgca cgtcgggcac tacgtgccag 1380
163   tatagcaacg actactactc gcaatgcctt tagagcgttg acttgctctt ggtctgtcca 1440
164   gacgggggca cgatagaatg cgggcacgca gggagctcgt agacattggg cttaatatat 1500
165   aagacatgct atgttgtatc tacattagca aatgacaaac aaatgaaaaa gaacttatca 1560
166   agcaaaaaaa aaaaaaaa aaaaaaaa          1588
169 <210> SEQ ID NO: 7
170 <211> LENGTH: 745
171 <212> TYPE: DNA
172 <213> ORGANISM: Trichoderma reesei
174 <220> FEATURE:
175 <221> NAME/KEY: terminator
176 <222> LOCATION: (1)..(745)
177 <223> OTHER INFORMATION: T. reesei cbh1 terminator
179 <400> SEQUENCE: 7
180   ggacctacc agtctcacta cggccagtgc ggcggtattg gctacagcgg cccacggtc 60
181   tgcgccagcg gcacaacttg ccaggctcctg aacccttact actctcagt cctgtaaagc 120

```

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/936,823

DATE: 10/04/2001
 TIME: 11:12:32

Input Set : A:\36838SEQ.txt
 Output Set: N:\CRF3\10042001\I936823.raw

```

182   tccgtgcgaa agcctgacgc accggtagat tcttggtgag cccgtatcat gacggcggcg 180
183   ggagctacat ggccccgggt gatttatattt ttttgtatct acttctgacc cttttcaaatt 240
184   atacggtcaa ctcatctttc actggagatg cggcctgctt ggtattgcga tgttgtcagc 300
185   ttggcaaatt gtggctttcg aaaacacaaa acgattcctt agtagccatg cattttaaga 360
186   taacggaata gaagaaagag gaaattaaaa aaaaaaaaaa aacaaacatc ccgttcataa 420
187   cccgtagaat cgccgctctt cgtgtatccc agtaccacgt caaagggtatt catgatcggt 480
188   caatgttgat attgttccgc cagtattggt ccacccccat ctccgcgaat ctctctttct 540
189   cgaacgcggt agtggctgct gccaatgggt aatgaccata gggagacaaa cagcataata 600
190   gcaacagtgg aaattagtgg cgcaataatt gagaacacag tgagaccata gctggcggcc 660
191   tggaaagcac tgttggagac caacttgtcc gttgcgaggc caacttgcatt tgctgtcaag 720
192   acgatgacaa cgtagccgag gaccc                                     745
195 <210> SEQ ID NO: 8
196 <211> LENGTH: 10
197 <212> TYPE: DNA
198 <213> ORGANISM: Artificial Sequence
200 <220> FEATURE:
201 <223> OTHER INFORMATION: Description of Artificial Sequence: annealed primer
203 <400> SEQUENCE: 8
204   taaccgcggt                                     10
207 <210> SEQ ID NO: 9
208 <211> LENGTH: 16
209 <212> TYPE: DNA
210 <213> ORGANISM: Artificial Sequence
212 <220> FEATURE:
213 <223> OTHER INFORMATION: Description of Artificial Sequence: annealed primer
215 <400> SEQUENCE: 9
216   ctagaccgcg gttaat                                     16
219 <210> SEQ ID NO: 10
220 <211> LENGTH: 1232
221 <212> TYPE: DNA
222 <213> ORGANISM: Trichoderma reesei
224 <220> FEATURE:
225 <221> NAME/KEY: promoter
226 <222> LOCATION: (1)..(1232)
227 <223> OTHER INFORMATION: T. reesei gpd1 promotor
229 <400> SEQUENCE: 10
230   gtcgacacga tatacaggcg cggctgatga taatgatgat cgagcatgac ttgatgctgt 60
231   atgtgacaat attgactgcg aggaaccatc aggtgtgtat ggatggaatc attctgtaac 120
232   caccaagggtg catgcatcat aaggattctc ctacgctcac caacaacgaa cgatggccat 180
233   gttagtgaag gcaccgtgat ggcaagatag aaccactatt gcattctgcg tccccacgca 240
234   cagtacgtca agtaacgtca aagccgccct cccgtaacct cgcccgttgt tgctcccccc 300
235   gattgcctca atcacatagt acctacctat gcattatggg cgccctcaac ccaccccccc 360
236   agattgagag ctaccttaca tcaatatggc cagcacctct tcggcgatac atactcgcca 420
237   ccccagccgg cgcgattgtg tgtactaggt aggtctgtac tataccagca ggagagggtg 480
238   tgcttggcaa tcgtgctcag ctgttaggtt gtacttgtat ggtacttgta aggtgggtcat 540
239   gcagttgcta aggtacctag ggagggattc aacgagccct gcttccaatg tccatctgga 600
240   taggatggcg gctggcgggg ccgaagctgg gaactcgcca acagtcatat gtaatagctc 660
241   aagttgatga taccgttttg ccagattaga tgcgagaagc agcatgaatg tcgctcatcc 720
242   gatgccgcatt caccgttgtg tcagaaacga ccaagctaag caactaaggt accttaccgt 780

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/936,823

DATE: 10/04/2001

TIME: 11:12:32

Input Set : A:\36838SEQ.txt

Output Set: N:\CRF3\10042001\I936823.raw

```

243 ccactatctc aggtaaccag gtactaccag ctaccctacc tgccgtgcct acctgcttta 840
244 gtgttaatat ttccacctcc ctctcaatc ttcttttccc tctctctctc tttttttttt 900
245 ctctctctct ttctttctcc taaccattcc taacaacatc gacattctct cctaataacc 960
246 agcctcgcaa atcctcagtt tgtatgtacg tacgtactac aatcatcacc acgatcgtcc 1020
247 gcccagacgat gcggcttctg ttgcctgcc cctcctctca ctcggtgcct tgacgagcta 1080
248 gcccgcagag gactctcctg cgtcaccaat ttttttccct atttaccctt cctccctctc 1140
249 tccctctcgt ttcttctctaa caaacaacca ccaccaaata ctctttggaa gctcacgact 1200
250 cagcagctc aattcgcaga tacaatatga ga 1232

```

253 <210> SEQ ID NO: 11

254 <211> LENGTH: 1129

255 <212> TYPE: DNA

256 <213> ORGANISM: Trichoderma reesei

258 <220> FEATURE:

259 <221> NAME/KEY: terminator

260 <222> LOCATION: (1)..(1129)

261 <223> OTHER INFORMATION: T. reesei gpd1 terminator

263 <400> SEQUENCE: 11

```

264 ggatcccgag cattgtctat gaatgcaaac aaaaatagta aataaatagt aattctggcc 60
265 atgacgaata gagccaatct gctccacttg actatcttgt gactgtatcg tatgtcgaac 120
266 ccttgactgc ccattcaaac aattgtaaag gaatatagct acaagttatg tctcacgttt 180
267 gcgtgcgagc ccgtttgtac gttattttga gaaagcgttg ccatacatg ctcacagtca 240
268 cttggcttac gatcatgttt gcgatcttcg gtaagaatac acagagtaac gattatctcc 300
269 atcgcttcta tgattaggta ctacagacaac acatgggaaa caagataacc atcgcatgca 360
270 aggtcgattc caatcatgat ctggactggg gtattccatc taagccatag taccctcgag 420
271 agaaggaatg gtaggacctc tcaggcgctc accatctgtg ctgcaaatac aagaaacccc 480
272 ccaaaagcac ctacctatct acctagagta actgcacgag aaaagaaaag gagcagaaga 540
273 agaatgatct caagaggccg tgaacgcaga aacacactcc tcccaacttt tcaagttttg 600
274 aacaaaaaaa gaaagatgag gactagaaga tggagtattt cttctctaga gagctctcgg 660
275 tgaggtgacc tgtcagggtt taccgcaaac cgtcgggtgt tctatccaat taatcaagtc 720
276 ccgcgcctcg cctcttctct cctgtccttt catagaatcc cgtctccttg ttgcttgatc 780
277 gaagcggggt tatcgacgcc accaaagatc ttgtcttggt gacttatcaa tccttttggtg 840
278 atcaaacagc ccccgagtga tcagatccgt aaaagaagaa gaagagtacg atttaaccag 900
279 accgaggaac aataaagcga gtaaataaca tcaaaataag agtctcgttg aaaattactt 960
280 gttcctcaat caatcccaac ccccttaaaa gcccttcccc ccatggtata tcccggcagt 1020
281 aggagagaga tatttccact accgctcacc accaagttag gcttgccgag agaagaggat 1080
282 gaatcagaag tgacaacaac gggttgagca catgggatat cggcgcgcc 1129

```

285 <210> SEQ ID NO: 12

286 <211> LENGTH: 5733

287 <212> TYPE: DNA

288 <213> ORGANISM: Aspergillus nidulans

290 <220> FEATURE:

291 <223> OTHER INFORMATION: (1-5733) Sequence of plasmid pAN52-1

293 <220> FEATURE:

294 <221> NAME/KEY: promoter

295 <222> LOCATION: (1)..(2129)

296 <223> OTHER INFORMATION: A. nidulans gpdA promoter

298 <220> FEATURE:

299 <221> NAME/KEY: gene

300 <222> LOCATION: (2130)..(2304)

Use of n and / or Xaa has been detected in the Sequence Listing. Review the Sequence Listing to ensure a corresponding explanation is present in the <220> to <223> fields of each sequence using n or Xaa.

VERIFICATION SUMMARY

DATE: 10/04/2001

PATENT APPLICATION: US/09/936,823

TIME: 11:12:33

Input Set : A:\36838SEQ.txt

Output Set: N:\CRF3\10042001\I936823.raw

L:9*M:270 C: Current Application Number differs, Replaced Application Number

L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:356 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12

L:360 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12

L:361 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12